

# Structure Learning and Statistical Estimation in Distribution Networks - Part I

Deepjyoti Deka\*, Scott Backhaus†, and Michael Chertkov‡

\*Corresponding Author. Electrical & Computer Engineering, University of Texas at Austin

†MPA Division, Los Alamos National Lab

‡Theory Division and the Center for Nonlinear Systems, Los Alamos National Lab

**Abstract**—Traditionally power distribution networks are either not observable or only partially observable. This complicates development and implementation of new smart grid technologies, such as those related to demand response, outage detection and management, and improved load-monitoring. In this two part paper, inspired by proliferation of metering technology, we discuss estimation problems in structurally loopy but operationally radial distribution grids from measurements, e.g. voltage data, which are either already available or can be made available with a relatively minor investment. In Part I, the objective is to learn the operational layout of the grid. Part II of this paper presents algorithms that estimate load statistics or line parameters in addition to learning the grid structure. Further, Part II discusses the problem of structure estimation for systems with incomplete measurement sets. Our newly suggested algorithms apply to a wide range of realistic scenarios. The algorithms are also computationally efficient – polynomial in time – which is proven theoretically and illustrated computationally on a number of test cases. The technique developed can be applied to detect line failures in real time as well as to understand the scope of possible adversarial attacks on the grid.

**Index Terms**—Power Distribution Networks, Power Flows, Structure/graph Learning, Voltage measurements, Transmission Lines.

## I. INTRODUCTION

The power grid is composed of a network of transmission and distribution lines that enable the transfer of electrical power from generators to loads. The design, operation and control of these networks is typically hierarchical with a major division occurring between the transmission network of high voltage lines connecting sub-stations and power plants, and the distribution network of medium and low voltage lines that connect the transmission sub-stations to the end-users. Here, we focus on distribution networks.

The design of distribution networks may appear to be loopy or meshed, however for practical engineering concerns, the vast majority of distribution grids are operated as "radial" networks, i.e. as a set of non-overlapping trees. Switches in the network are used to achieve one radial configuration out of many possibilities. Each tree in the network has a

substation at the root and customers positioned at the other nodes. Switching from one tree-like operational configuration to another is typically caused by system upsets, e.g. faults and outages, and may occur few times a day or even an hour.

The radial configuration distinguishes distribution networks from transmission networks that generally have multiple loops energized all the time to guarantee continuous delivery of power to every node, even in case of occasional line faults and outages. Radial configurations and one-way flow of power have led to much less monitoring, observability, and state estimation in distribution as compared to meshed transmission networks [1]. The recent proliferation of smart grid technology, including smart meters that measure electricity consumption at the node level, is creating a new opportunities to extract information important to grid operators and planners. Such efforts are also getting additional attention in view of mounting concerns over data security and protection of user privacy [2].

In this paper (Part I), we seek to *develop low-complexity algorithms to learn the current operational structure in 'radial' distribution networks using only nodal measurements*. Nodal measurements may include voltage magnitudes, voltage phase (potentially), and power injections and are typically available at smart meters, pole-mount or pad-mount transformers, and distribution phasor measurement units (PMU). Accurate structural estimation impacts many important applications including failure identification [3], outage management, and recovery following major and minor disruptions (e.g. hurricanes to individual lightning strikes), grid reconfiguration [4] for power flow optimization and generation scheduling [1], [4]–[6], and quantifying the need for additional meter placement. From an adversarial viewpoint, our work can be viewed as low-intrusion learning by a rogue agent interested in estimating the grid structure for a data attack [7], [8]. In the subsequent part (Part II), we will look at developing algorithms that are able to estimate the statistics of power consumption at the grid nodes or estimate the parameters of operational lines in addition to determining the grid's radial structure. Further, we will analyze learning the operational grid structure with missing data, where observations from a subset of nodes are not available.

## A. Related Work

Our work falls in the broad category of 'graph learning' problems that have been approached from different directions.

D. Deka is with the Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712. Email: deepjyotideka@utexas.edu

S. Backhaus is with the MPA Division of LANL, Los Alamos, NM 87544. Email: \*backhaus@lanl.gov

M. Chertkov is with the Theory Division and the Center for Nonlinear Systems of LANL, Los Alamos, NM 87544. Email: \*chertkov@lanl.gov

For general graphs and graphical models [9], maximum-likelihood structure estimation has been researched in several papers by utilizing prior information such as sparsity of the parameter space [10], [11], size of the graph neighborhood [12], etc. Techniques employed include both traditional convex optimization [10], [13] as well as greedy learning [12]. For power grids, structure estimation techniques discussed in the literature can be classified based on the type of measurements available as well as assumptions made regarding grid structure and user behavior. In [14], a maximum likelihood estimator (MLE) with regularizers for low-rank and sparsity is used to recover the grid structure using locational marginal prices (LMPs). In [15], a model using bus phase angles as a Markov random field for the DC power flow builds a dependency graph based approach to detect faults in grids. In work specific to radial distribution grids, [16] provides a structure identification algorithm that uses signs within the inverse covariance matrix (or concentration matrix) of voltage measurements to generate a minimum spanning tree. In [3], topology identification with limited measurements in a distribution grid with Gaussian loads is used to design a machine learning (ML) estimate with approximate schemes. Our work uses ordering of second moments, not a ML approach, to reconstruct a radial grid sequentially from the leaves to the root, making it distinct from previous work. Our algorithm design is based on a linear coupled approximation for lossless AC power flow that is idealized but practical [17], [18] for analyzing distribution grids where the line and voltage characteristics limit the accuracy of traditional approximations. Unlike related work, our topology learning algorithm is agnostic to the load profile distributions or variability in line impedances and requires only a less restrictive assumption on the correlation of load profiles.

The rest of the paper is organized as follows. We start the next section with a description of the distribution grid and summary of the learning problems discussed in the manuscript. Section III and Appendix A describe the linear coupled (LC) power flow model and its special case, the DC-resistive power flow model. Statistical trends in observed nodal measurements are discussed in Section III-C. Next, we use the derived results to design algorithms for learning the distribution grid structure using the power flow models in Section V. Simulation results elucidating the performance of our algorithms on test distribution grids are presented in Section VI. Finally, Section VII concludes and suggests future directions that will be explored in a subsequent work.

## II. TECHNICAL PRELIMINARIES

The structure of a radial distribution network has important features that motivates our algorithm development. We discuss the radial structure in detail here and introduce the notation used in this paper. We then formulate the learning problem tackled in this paper in terms of its input data and deliverables and discuss the underlying motivation.

### A. Structure of Radial Distribution Network

We consider a meshed distribution network which is operated as a union of non-intersecting ‘radial’ trees, i.e. a

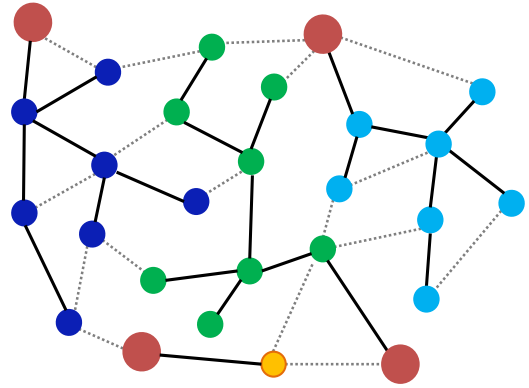


Fig. 1. Schematic layout of an example of a distribution grid fed from 4 substations, with substations represented by large red nodes. The operational grid has a base-4 spanning forest configuration that is formed by solid lines (black). Dotted grey lines represent open switches. Each tree in the forest has one substation at the root. The load nodes within each tree are marked with the same color.

spanning forest, by configuring switches as shown in Fig. 1. There are exponentially many (in the number of switches) possible configurations of spanning forests. The grid-graph with all the switches closed is denoted  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , where  $\mathcal{V}$  is the set of nodes of the graph and  $\mathcal{E}$  is the set of undirected edges of the graph. We denote nodes with single Roman letter subscripts  $a$  and undirected edges with pairs of Roman letter subscripts  $(ab)$ . The operational grid is a forest denoted by  $\mathcal{F}$  which spans all the nodes in  $\mathcal{V}$ . Specifically,  $\mathcal{F}$  is a special subgraph of  $\mathcal{G}$  ( $\mathcal{F} \subset \mathcal{G}$ ) such that

- $\mathcal{F}$  is a union of  $K$  non-overlapping trees covering all the nodes of the graph
- Each tree contains exactly one of the  $K$  bases (substations),  $\mathcal{F} = \cup_{k=1, \dots, K} \mathcal{T}_k$ .

The distribution system  $\mathcal{F}$  is a ‘**base-constrained spanning forest**’ with operational edges  $\mathcal{E}^{\mathcal{F}}$  where  $\mathcal{E}^{\mathcal{F}} \subset \mathcal{E}$ . Table I provides other relevant notations (nomenclature) used throughout this manuscript to denote various nodal and edge features of the grid  $\mathcal{G}$  and the operational forest  $\mathcal{F}$ .

### B. Problem Formulation and Contribution

We consider large distribution grids where the utility (observer) is unsure of the grid configuration because of insufficient or inaccurate switching data, perhaps caused by a recent system upset. Alternatively, we could take the point of view of a third party observer, who may be an aggregator or adversary, trying to extract the current forest configuration from available nodal measurements. We assume the current spanning forest configuration is kept intact sufficiently long for load profiles at grid nodes to attain a steady distribution (longer than the fluctuations but shorter than changes in the mean load).

We assume that the observer has access to nodal measurements, but not edge measurements – an assumption consistent with the recent expansion of smart grid monitoring devices. Smart meters generally provide nodal voltages and power injections at fine spatial resolution, i.e. at the individual customer level, but they do not provide any edge flow data.

TABLE I  
NOTATION TABLE

$\mathcal{F} \in \mathcal{G}$	a particular forest configuration of the physical distribution network $\mathcal{G}$
$\mathcal{V}$	vertex set of $\mathcal{G}$
$N$	# of nodes other than sub-stations in $\mathcal{V}$
$K$	number of sub-stations in the network
$\mathcal{E}$	edge set of $\mathcal{G}$
$\mathcal{E}^{\mathcal{F}}$	set of edges operational within $\mathcal{F}$
$\mathcal{T}_k \in \mathcal{F}$	tree within $\mathcal{F}$ containing the $k^{th}$ sub-station
$M_k$	reduced incidence matrix of the tree $\mathcal{T}_k$
$\mathcal{E}^{\mathcal{T}_k}$	set of edges in $\mathcal{T}_k$
$\mathcal{V}^{\mathcal{T}_k}$	set of nodes in $\mathcal{T}_k$
path between $a, b \in \mathcal{V}^{\mathcal{T}_k}$	subset of edges from $\mathcal{E}^{\mathcal{T}_k}$ s.t. each node with edge in the subset, except $a$ and $b$ , contributes exactly two edges
$\mathcal{E}_a^{\mathcal{T}_k}$	path from $a$ to slack bus in $\mathcal{T}_k$
$b$ is a descendant of $a$	$a$ contributes $\mathcal{E}_b^{\mathcal{T}_k}$
$b$ is the parent of $a$	$(ab) \in \mathcal{E}_a^{\mathcal{T}_k}$ and $b$ contributes $\mathcal{E}_a^{\mathcal{T}_k}$
$\theta_a, v_a$	voltage phase and magnitude resp. at bus $a$
$\theta, v$	Vector of non-substation voltage phases and magnitudes resp.
$\epsilon_a$	$= 1 - v_a$ , voltage deviation at node $a$
$\epsilon$	$= 1 - v$ , voltage deviation
$p_a, q_a$	resp. active and reactive power injection/consumption (+/-) at bus $a$
$p, q$	Vector of non-substation power injections/consumptions
$\beta_{ab}, g_{ab}, r_{ab}, x_{ab}$	susceptance, conductance, resistance, reactance resp. of edge $(ab)$
$\beta, g, r, x$	diagonal matrix of line susceptances, conductances, resistances, reactances resp.
$\Sigma_y$	matrix of second moments for variable $y$
$\Omega_y$	matrix of covariances for variable $y$
$\mu_y$	vector of means for variable $y$
$H_y$	reduced weighted Laplacian matrix with edge weights in $y$
$D_a^{\mathcal{T}_k}$	Descendants (including itself) of $a$ in $\mathcal{T}_k$
$\mathcal{M}$	set of unobserved nodes

Additional instrumentation is emerging for pole-mount or pad-mount transformers [19], however, these new devices still only provide nodal voltages and aggregated customer power injections.<sup>1</sup> Some edge flow data is available to utilities, however, this is generally at a few select locations in the distribution grid, e.g. at the substation/root node, voltage regulators, reclosers, or other major utility equipment. These select locations may also have nodal and edge data from another emerging technology, i.e. distribution grid PMUs [20]. However, we continue to restrict our input data to nodal values, which is consistent with the new, ubiquitous sensing provided by smart meters.

The nodal devices provide the observer with temporal samples of the nodal voltage magnitudes. The observer seeks to use these samples to learn the current configuration of switches that determine the ‘base-constrained spanning forest’. To supplement the voltage magnitude samples, the observer has historical information about statistics of the nodal consumption.

<sup>1</sup>We use the term ‘power injection’, ‘power consumption’ and ‘load’ interchangeably to denote the power profile at each interior (non-substation) node of the distribution system.

### III. POWER FLOW MODELS AND STATISTICAL CORRELATIONS

Our approach to the structure learning problem relies on linearized PF models on radial spanning forests that enable efficient reconstruction of the grid structure via a second-moment analysis. The most general of the two, termed the Linear Coupling (LC) model, ignores losses of active and reactive powers and consistently assumes small voltage magnitude and phase drops between connected nodes. For tree-like distribution grids, the LC-PF model becomes equivalent to the LinDistFlow PF model in [17], [18]. The second model considered in the paper, coined the DC-resistive model, corresponds to the special resistance dominating case of the LC-PF model. These PF models are described in more detail in Appendix A.

#### A. Linear Coupled Power Flow (LC-PF) model

As noted in Eqs. (25,26) in Appendix A, the LC-PF model is derived from the general AC power flow model by assuming small voltage magnitude deviations and phase differences between neighboring buses in the grid. It is convenient to restate the linear equations in LC-PF model in matrix form as:

$$p = H_g \epsilon + H_\beta \theta, \quad q = H_\beta \epsilon - H_g \theta \quad (1)$$

where  $p, q, \epsilon$  and  $\theta$  are defined in Table I.  $H_g$  and  $H_\beta$  are the weighted graph Laplacian matrices associated with forest  $\mathcal{F}$  such that

$$H_g(a, b) = \begin{cases} \sum_{c: (a,c) \in \mathcal{E}^{\mathcal{F}}} g_{ac} & \text{if } b = a \\ -g_{ab} & \text{if } (ab) \in \mathcal{E}^{\mathcal{F}} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

$H_\beta$  has a similar structure with  $g$ -weights replaced by  $\beta$ -weights. The weighted graph Laplacians can be stated in terms of the directed incidence matrix  $M$  as

$$H_g = M^T g^{\mathcal{F}} M, \quad H_\beta = M^T \beta^{\mathcal{F}} M \quad (3)$$

Here,  $g^{\mathcal{F}}$  and  $\beta^{\mathcal{F}}$  are diagonal matrices representing, respectively, line conductances and susceptances for edges within  $\mathcal{F}$ .  $M$  is the edge to node directed incidence matrix of  $\mathcal{F}$ . See Fig. 2 for an example. Every row  $m_{ab}$  in  $M$  is equal to  $\pm(e_a^T - e_b^T)$  and represents the directed edge  $(ab)$ , where the direction of an edge is chosen arbitrarily.  $e_a \in \mathbb{R}^{N+K}$  is the standard basis vector associated with the vertex  $a$ , with 1 at the  $a^{th}$  position and zero everywhere else. We can combine Eqs. (1) and express the complex power flows as:

$$p + iq = M^T (g^{\mathcal{F}} + i\beta^{\mathcal{F}}) M (\epsilon - i\theta) \quad (4)$$

Both  $H_\beta$  and  $H_g$  are weighted graph Laplacians and are degenerate — showing  $K$  zero-eigenvalues associated with the freedom in fixing phase and voltage deviation (from nominal) at any node within each tree of the forest. It is natural to fix phases and voltages at the sub-stations making these ‘slack buses’  $a_k$  for trees  $\mathcal{T}_k$  of the (operational) forest,  $\mathcal{F}$  such that  $\theta_{a_k} = \epsilon_{a_k} = 0$ , for any  $k$ ,  $1 \leq k \leq K$ . Formally, elimination of the set of  $K$  sub-stations corresponds to elimination of  $K$

components from all the vectors contributing Eqs. (1), and reduction of  $K$  rows and  $K$  columns from the weighted Laplacian matrices. All the eigenvalues of the resulting reduced graph Laplacian matrices are thus strictly positive.

Without loss of generality, we will use the same notation for the original and reduced dimension variables  $\theta, \varepsilon, p$  and  $q$  and also refer to Eqs. (1,4) as applied to the reduced vectors of dimension  $N \times 1$ . We will also keep notations,  $H_\beta$  and  $H_g$  for the reduced graph Laplacian matrices, and  $M$  for the reduced incidence matrix respectively. The reduced  $M$  has a block diagonal structure:  $M = \text{diag}(M_1, M_2, \dots, M_K)$ , where,  $M_k$  is the invertible reduced incidence matrix of tree  $\mathcal{T}_k$  in  $\mathcal{F}$ . Thus,  $M$  and correspondingly  $H_\beta$  and  $H_g$  are full rank, invertible and block-diagonal matrices. Inverting the linear non-degenerate Eqs. (1) we arrive at

$$\theta = M^{-1} x^\mathcal{F} M^{-1T} p - M^{-1} r^\mathcal{F} M^{-1T} q = H_{1/x}^{-1} p - H_{1/r}^{-1} q \quad (5)$$

$$\varepsilon = M^{-1} r^\mathcal{F} M^{-1T} p + M^{-1} x^\mathcal{F} M^{-1T} q = H_{1/r}^{-1} p + H_{1/x}^{-1} q \quad (6)$$

where  $H_{1/r} \doteq M^T r^{\mathcal{F}-1} M$  and  $H_{1/x} \doteq M^T x^{\mathcal{F}-1} M$ .  $r^\mathcal{F}$  and  $x^\mathcal{F}$  are diagonal matrices representing, respectively, line resistances and reactances within the forest  $\mathcal{F}$ . Their relation to  $g^\mathcal{F}$  and  $\beta^\mathcal{F}$  are expressed in Eqs. (28).

### B. Relations between second moments

The real and reactive nodal power injections  $p$  and  $q$  in Eqs. (5,6) fluctuate because of exogenous processes, and their second moments are related by:

$$\begin{aligned} \mathbb{E}[\theta\theta^T] &= H_{1/x}^{-1} \mathbb{E}[pp^T] H_{1/x}^{-1} + H_{1/r}^{-1} \mathbb{E}[qq^T] H_{1/r}^{-1} \\ &\quad - H_{1/x}^{-1} \mathbb{E}[pq^T] H_{1/r}^{-1} - H_{1/r}^{-1} \mathbb{E}[qp^T] H_{1/x}^{-1} \\ \Rightarrow \Sigma_\theta &= H_{1/x}^{-1} \Sigma_p H_{1/x}^{-1} + H_{1/r}^{-1} \Sigma_q H_{1/r}^{-1} \\ &\quad - H_{1/x}^{-1} \Sigma_{pq} H_{1/r}^{-1} - \left[ H_{1/x}^{-1} \Sigma_{pq} H_{1/r}^{-1} \right]^T \end{aligned} \quad (7)$$

$$\begin{aligned} \text{Similarly, } \Sigma_\varepsilon &= H_{1/r}^{-1} \Sigma_p H_{1/r}^{-1} + H_{1/x}^{-1} \Sigma_q H_{1/x}^{-1} \\ &\quad + H_{1/r}^{-1} \Sigma_{pq} H_{1/x}^{-1} + \left[ H_{1/r}^{-1} \Sigma_{pq} H_{1/x}^{-1} \right]^T \end{aligned} \quad (8)$$

$$\begin{aligned} \Sigma_{\theta\varepsilon} &= H_{1/x}^{-1} \Sigma_p H_{1/r}^{-1} - H_{1/r}^{-1} \Sigma_q H_{1/x}^{-1} \\ &\quad + H_{1/x}^{-1} \Sigma_{pq} H_{1/x}^{-1} - H_{1/r}^{-1} \Sigma_{qp} H_{1/r}^{-1} \end{aligned} \quad (9)$$

These formulas are the basis for reconstruction/learning analysis in the rest of the paper.

### C. DC-resistive model

The DC-resistive PF model (see Appendix A) is an extremal case of the LC-PF model realized when line reactance can be ignored in comparison with resistance ( $x/r \rightarrow 0$ ). The relation between the statistics of active powers and voltage second-order moments deviations reduces to

$$\mathbb{E}[pp^T] = H_g \mathbb{E}[\varepsilon\varepsilon^T] H_g \Rightarrow \Sigma_\varepsilon = H_g^{-1} \Sigma_p H_g^{-1} \quad (10)$$

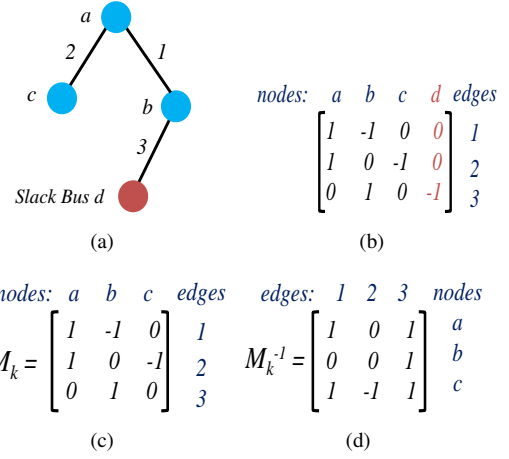


Fig. 2. Illustration of the reduced incidence matrix construction on the example of a tree with 4 nodes. (a) Tree graph with four nodes ( $a, b, c, d$ ) and three edges (1, 2, 3). (b) Complete directed incidence matrix. The “directed” freedom in choosing edge orientations is fixed as follows, 1 = ( $ab$ ), 2 = ( $ac$ ), 3 = ( $bd$ ). (c) The column corresponding to the slack bus (bus  $d$ ) is removed in the reduced incidence matrix  $M_k$ . (d) Inverse of the reduced incidence matrix  $M_k$ .

## IV. TRENDS IN SECOND MOMENTS OVER TREE NETWORKS

We now derive key results related to the second moments in voltage magnitudes that arise from the properties of the forest  $\mathcal{F}$ . We denote the unique path from node  $a$  to the slack bus in tree  $\mathcal{T}_k$  by  $\mathcal{E}_a^{\mathcal{T}_k}$ . From [21], the inverse of the reduced incidence matrix of a tree has the following special structure:

$$M_k^{-1}(a, r) = \begin{cases} 1 & \text{if edge } r \in \mathcal{E}_a^{\mathcal{T}_k} \text{ is directed} \\ & \text{along path from } a \text{ to slack bus,} \\ -1 & \text{if edge } r \in \mathcal{E}_a^{\mathcal{T}_k} \text{ is directed} \\ & \text{against path from } a \text{ to slack bus,} \\ 0 & \text{if edge } r \notin \mathcal{E}_a^{\mathcal{T}_k} \end{cases} \quad (11)$$

Here, the direction of edge  $r = (cd)$  is specified by its representative row  $m_{cd}$  in the directed incidence matrix. For example, if  $m_{cd} = e_c^T - e_d^T$ , the edge is directed from node  $c$  to node  $d$ , whereas for  $m_{cd} = e_d^T - e_c^T$ , the direction is from node  $d$  to node  $c$ .

An immediate corollary of (11) is that  $M^{-1}(a, r) = 0$  if edge  $r$  and node  $a$  lie on separate trees within the forest  $\mathcal{F}$ , a fact consistent with the block diagonal structure of  $M$ . Using (11) in  $H_g^{-1} = M^{-1} g^{\mathcal{F}-1} M^{-1T}$ , we derive for forest  $\mathcal{F}$

$$H_g^{-1}(a, b) = 0 \text{ if } a, b \text{ are on different trees } \mathcal{T}_k \text{ and} \quad (12)$$

$$\begin{aligned} H_g^{-1}(a, b) &= \sum_{r \in \mathcal{E}^{\mathcal{T}_k}} M^{-1}(a, r) g^{\mathcal{F}-1}(r, r) M^{-1}(b, r) \text{ if } a, b \in \mathcal{T}_k \\ &= \sum_{(cd) \in \mathcal{E}_a^{\mathcal{T}_k} \cap \mathcal{E}_b^{\mathcal{T}_k}} \frac{1}{g_{cd}} \text{ if } a, b \in \mathcal{T}_k \end{aligned} \quad (13)$$

Thus,  $H_g^{-1}(a, b)$  is equal to the sum of the inverse conductances of lines that are common to the paths from both nodes to the slack bus. If no such line exists, the corresponding entry in  $H_g^{-1}$  is 0. See Fig. 3(a) for illustration. Similar results hold

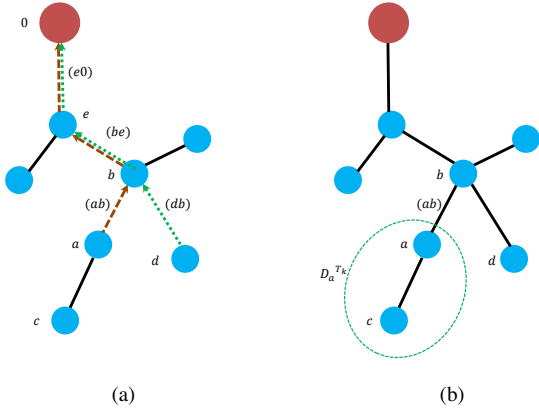


Fig. 3. Schematic layout of a distribution grid tree  $\mathcal{T}_k$ . The sub-station node represented by large red node is the slack bus. (a) Dotted lines represent the paths from nodes  $a$  and  $d$  to the slack bus. The common edges on those paths are edges  $(be)$  and  $(e0)$ . Thus,  $H_g^{-1}(a, d) = 1/g_{be} + 1/g_{e0}$ . (b) the set of descendants of  $a$ , denoted by  $D_a^{\mathcal{T}_k}$ . Here, nodes  $a$  and  $c$  are descendants of both nodes  $a$  and  $b$ , while node  $d$  is not a descendant of  $a$  but only of node  $b$ .

for other measurement matrices like

$$H_{1/r}^{-1}(a, b) = \begin{cases} \sum_{(cd) \in \mathcal{E}_a^{\mathcal{T}_k} \cap \mathcal{E}_b^{\mathcal{T}_k}} r_{cd} & \text{if nodes } a, b \in \mathcal{T}_k \\ 0 & \text{otherwise,} \end{cases} \quad (14)$$

Let  $D_a^{\mathcal{T}_k}$  denote the set of descendants of node  $a$  within the tree  $\mathcal{T}_k$ . We call  $b$  a descendent of  $a$ , if  $a$  lies on the (unique) path from  $b$  to the slack bus of  $\mathcal{T}_k$ , also including  $a$  itself in the set of its descendants. We call  $b$  the parent of  $a$  within  $\mathcal{T}_k$  (there can only be one) if  $(ab) \in \mathcal{T}_k$  and  $a$  is an immediate descendant of  $b$  as illustrated in Fig. 3(b).

The following statement holds.

**Lemma 1.** For two nodes,  $a$  and its parent  $b$ , in tree  $\mathcal{T}_k$

$$H_g^{-1}(a, c) - H_g^{-1}(b, c) = \begin{cases} \frac{1}{g_{ab}} & \text{if node } c \in D_a^{\mathcal{T}_k} \\ 0 & \text{otherwise,} \end{cases} \quad (15)$$

*Proof:* For any node  $c$  which belongs to a tree not containing nodes  $a$  and  $b$ ,  $H_g^{-1}(a, c) - H_g^{-1}(b, c) = 0$  according to (12). Now, focus on nodes contained, together with  $a$  and  $b$ , within  $\mathcal{T}_k$ . Since  $b$  is  $a$ 's parent,  $\mathcal{E}_a^{\mathcal{T}_k} = \mathcal{E}_b^{\mathcal{T}_k} \cup \{(ab)\}$  and we derive (also validating on the illustrative example in Fig. 3(b)) that for any node  $c$  in tree  $\mathcal{T}_k$ ,

$$\begin{aligned} \mathcal{E}_a^{\mathcal{T}_k} \cap \mathcal{E}_c^{\mathcal{T}_k} &= \mathcal{E}_b^{\mathcal{T}_k} \cap \mathcal{E}_c^{\mathcal{T}_k} & \text{if node } c \notin D_a^{\mathcal{T}_k} \\ \mathcal{E}_a^{\mathcal{T}_k} \cap \mathcal{E}_c^{\mathcal{T}_k} &= [\mathcal{E}_b^{\mathcal{T}_k} \cap \mathcal{E}_c^{\mathcal{T}_k}] \cup \{(ab)\} & \text{if node } c \in D_a^{\mathcal{T}_k} \end{aligned}$$

resulting in Eq. (15). ■

We now prove our main results regarding trends in second moments of deviations in voltage magnitudes ( $\epsilon$ ) along any tree in the network. The results are conditioned on the following assumption regarding correlations in power injections at the non-substation buses.

**Assumption 1:** For any two buses  $a$  and  $b$  drawing power from the same distribution sub-station,  $\Sigma_p(a, b) > 0$ ,  $\Sigma_q(a, b) > 0$ ,  $\Sigma_{pq}(a, b) > 0$ .

Note that this assumption holds, in particular, if the overall node balance is such that each non-substation node  $a$  always

consumes strictly more than it produces in active and reactive powers, i.e.  $p_a < 0, q_a < 0$ . This assumption is certainly true in any distribution grid with small and/or moderate penetration of renewables [22]. However, the assumption is also reasonable for a system with significant penetration of generation which is still dominated in average by the consumption.

**Theorem 1.** If node  $a \neq b$  is a descendant of node  $b$  within tree  $\mathcal{T}_k$ , then for the DC-resistive model,  $\Sigma_\epsilon(a, a) > \Sigma_\epsilon(b, b)$ .

*Proof:* We first show that for any node  $a$  and its parent  $b$ ,  $\Sigma_\epsilon(a, a) > \Sigma_\epsilon(b, b)$ . Consider  $\Sigma_\epsilon(a, a) - \Sigma_\epsilon(a, b)$ . From Eq. (10), we derive

$$\begin{aligned} \Sigma_\epsilon(a, a) - \Sigma_\epsilon(a, b) &= \sum_{c, d} H_g^{-1}(a, c) \Sigma_p(c, d) \\ &\quad (H_g^{-1}(a, d) - H_g^{-1}(b, d)) \end{aligned} \quad (16)$$

$$\begin{aligned} & \text{(using Lemma 1)} = \sum_{c, d} H_g^{-1}(a, c) \Sigma_p(c, d) \frac{1}{g_{ab}} \mathbf{1}(d \in D_a^{\mathcal{T}_k}) \\ & \Rightarrow \Sigma_\epsilon(a, a) - \Sigma_\epsilon(a, b) > 0 \text{ (using Assumption 1)} \end{aligned} \quad (17)$$

$$\text{Also, } \Sigma_\epsilon(a, b) - \Sigma_\epsilon(b, b) = \sum_{c \in D_a^{\mathcal{T}_k}, d} H_g^{-1}(b, d) \Sigma_p(c, d) \frac{1}{g_{ab}} > 0 \quad (18)$$

Combining Eqs. (17) and (18) we derive  $\Sigma_\epsilon(a, a) > \Sigma_\epsilon(a, b) > \Sigma_\epsilon(b, b)$ . Since node  $a$  is a descendant of node  $b$ , there is a path  $a, c_1, \dots, c_r, b$ , such that each node in the path is a parent of its predecessor. Then, we derive  $\Sigma_\epsilon(a, a) > \Sigma_\epsilon(c_1, c_1) > \dots > \Sigma_\epsilon(c_r, c_r) > \Sigma_\epsilon(b, b)$ . ■

The following theorem is the LC-PF version of Theorem 1.

**Theorem 2.** If node  $a \neq b$  is a descendant of node  $b$  on tree  $\mathcal{T}_k$  in forest  $\mathcal{F}$ , then  $\Sigma_\epsilon(a, a) > \Sigma_\epsilon(b, b)$  in the LC-PF model.

*Proof:* Consider Eq. (8). Notice that the right hand side has four constituent terms ( $H_{1/r}^{-1} \Sigma_p H_{1/r}^{-1}$ ,  $H_{1/x}^{-1} \Sigma_q H_{1/x}^{-1}$ ,  $H_{1/r}^{-1} \Sigma_{pq} H_{1/x}^{-1}$  and  $H_{1/x}^{-1} \Sigma_{qp} H_{1/r}^{-1}$ ). We denote each of these terms by  $\Sigma_\epsilon^j$  where  $j \in \{1, 2, 3, 4\}$ . For each individual term, applying Assumption 1 and the analysis in Theorem 1, we find that  $\Sigma_\epsilon^j(a, a) > \Sigma_\epsilon^j(b, b)$  if node  $a$  is a descendant of node  $b$ , other than itself. Thus, the statement also holds for the sum. ■

We now focus on evaluating the term  $\mathbb{E}[(\epsilon_a - \epsilon_b)^2]$ , which is the expected value of the squared difference between two node voltage deviations ( $\epsilon$ ). For any two nodes  $a$  and  $b$ , the DC-resistive model yields:

$$\begin{aligned} \mathbb{E}[(\epsilon_a - \epsilon_b)^2] &= \Sigma_\epsilon(a, a) - \Sigma_\epsilon(a, b) - \Sigma_\epsilon(b, a) + \Sigma_\epsilon(b, b) \\ & \text{(using (16))} = \sum_{c, d} H_g^{-1}(a, c) \Sigma_p(c, d) (H_g^{-1}(a, d) - H_g^{-1}(b, d)) \\ & \quad - \sum_{c, d} H_g^{-1}(b, c) \Sigma_p(c, d) (H_g^{-1}(a, d) - H_g^{-1}(b, d)) \\ & = \sum_{c, d} (H_g^{-1}(a, c) - H_g^{-1}(b, c)) \Sigma_p(c, d) \\ & \quad (H_g^{-1}(a, d) - H_g^{-1}(b, d)) \end{aligned} \quad (19)$$

Our next Lemma follows directly by applying Lemma 1 to Eq. (19).

**Lemma 2.** For two nodes,  $a$  and its parent  $b$  belonging to tree  $\mathcal{T}_k$ , in the DC-resistive model, we derive  $\mathbb{E}[(\epsilon_a - \epsilon_b)^2] = \sum_{c, d \in D_a^{\mathcal{T}_k}} \frac{1}{g_{ab}} \Sigma_p(c, d)$

For the LC-PF model, we evaluate the expression  $\mathbb{E}[(\epsilon_a - \epsilon_b)^2]$  as well. In this case, for two nodes,  $a$  and its parent  $b$ , that lie on tree  $\mathcal{T}_k$ , we arrive at

$$\mathbb{E}[(\epsilon_a - \epsilon_b)^2] = \Sigma_\epsilon(a, a) - \Sigma_\epsilon(a, b) - \Sigma_\epsilon(b, a) + \Sigma_\epsilon(b, b)$$

where  $\Sigma_\epsilon$  is given by Eq. (8). Let  $\Sigma_\epsilon^1$  and  $\Sigma_\epsilon^2$  represent the symmetric terms  $H_{1/r}^{-1}\Sigma_p H_{1/r}^{-1}$  and  $H_{1/x}^{-1}\Sigma_q H_{1/x}^{-1}$  respectively in  $\Sigma_\epsilon$ . Extending the result of Lemma 2, we derive

$$\Sigma_\epsilon^1(a, a) - \Sigma_\epsilon^1(a, b) - \Sigma_\epsilon^1(b, a) + \Sigma_\epsilon^1(b, b) = \sum_{c, d \in D_a^{\mathcal{T}_k}} r_{ab}^2 \Sigma_p(c, d) \quad (20)$$

$$\Sigma_\epsilon^2(a, a) - \Sigma_\epsilon^2(a, b) - \Sigma_\epsilon^2(b, a) + \Sigma_\epsilon^2(b, b) = \sum_{c, d \in D_a^{\mathcal{T}_k}} x_{ab}^2 \Sigma_q(c, d) \quad (21)$$

Similarly, let  $\Sigma_\epsilon^3 = H_{1/r}^{-1}\Sigma_{pq} H_{1/x}^{-1}$  and  $\Sigma_\epsilon^4 = \Sigma_\epsilon^{3T} = [H_{1/r}^{-1}\Sigma_{pq} H_{1/x}^{-1}]^T$ , the non-symmetric terms in  $\Sigma_\epsilon$ . Using Lemma 1 with Eq. (19), we get

$$\begin{aligned} \Sigma_\epsilon^3(a, a) - \Sigma_\epsilon^3(a, b) - \Sigma_\epsilon^3(b, a) + \Sigma_\epsilon^3(b, b) &= \sum_{c, d \in D_a^{\mathcal{T}_k}} r_{ab} \Sigma_{pq}(c, d) x_{ab} \\ &= \Sigma_\epsilon^4(a, a) - \Sigma_\epsilon^4(a, b) - \Sigma_\epsilon^4(b, a) + \Sigma_\epsilon^4(b, b) \end{aligned} \quad (22)$$

Combining Eqs. (20,21,22) we arrive at the following Lemma.

**Lemma 3.** *In the LC-PF model,  $\mathbb{E}[(\epsilon_a - \epsilon_b)^2] = \sum_{c, d \in D_a^{\mathcal{T}_k}} r_{ab}^2 \Sigma_p(c, d) + x_{ab}^2 \Sigma_q(c, d) + 2r_{ab}x_{ab}\Sigma_{pq}(c, d)$ , holds for a node  $a$  and its parent  $b$  belonging to the (operational) tree  $\mathcal{T}_k$ .*

## V. LEARNING STRUCTURE OF BASE-CONSTRAINED SPANNING FOREST

Here, we propose Algorithm 1 to learn the structure of the distribution network using properties of voltage deviations for the LC-PF model. The polynomial time algorithm, based on the Theorems proved in the previous section, requires positivity of the correlation between nodal power injections (Assumption 1). The Algorithm is also agnostic to the probability distribution of active and reactive power injections. To reconstruct the Base-Constrained Spanning Forest ( $\mathcal{F} = \cup_{k=1, \dots, K} \mathcal{T}_k$ ), Algorithm 1 takes as input ‘ $m$ ’ measurements of nodal voltage deviations. These measurements are used to create the empirical voltage deviation second moment matrix  $\Sigma_\epsilon$ . The voltage deviations  $\epsilon_a^j$  at the substation nodes are assumed to be zero which implies the elements of the row  $\Sigma_\epsilon(a, :) = 0$  for each of the  $K$  substation nodes. The observer has prior information (or estimates using power injection measurements) for the true second moment matrix of power injections  $\Sigma_p$  for non-substation nodes.

**Algorithm 1 Overview:** We reconstruct each tree within the distribution grid forest sequentially moving from the leaves to the root nodes. At every stage,  $U$  represents the set of undiscovered nodes that are not part of the current reconstructed tree while  $L$  represents the set of ‘current leaves’ (nodes that are in the current reconstructed tree but with undiscovered parents). At each iteration, Step 4 selects the node  $b^*$  from set  $U$  with the largest second moment of voltage deviation. Next

## Algorithm 1 Base Constrained Spanning Forest Learning: LC-PF Model

**Input:** True  $\Sigma_p, \Sigma_q$  and  $\Sigma_{pq}$ ,  $m$  voltage deviation observations  $\epsilon^j, 1 \leq j \leq m$ , all line resistances  $r$  and reactances  $x$

- 1: Compute  $\Sigma_\epsilon(a, a) = \sum_{j=1}^m \epsilon_a^j \epsilon_a^j / m$  for all nodes  $a$ .
- 2: Undiscovered Set  $U \leftarrow \{1, 2, \dots, N + K\}$ , Leaf Set  $L \leftarrow \emptyset$ , Descendant Sets  $D_a \leftarrow \{a\} \forall$  nodes  $a$ .
- 3: **while** ( $U \neq \emptyset$ ) **do**
- 4:    $b^* \leftarrow \max_{b \in U} \Sigma_\epsilon(b, b)$
- 5:   **for all**  $a \in L$  **do**
- 6:     **if**  $\sum_{j=1}^m (\epsilon_a^j - \epsilon_{b^*}^j)^2 / m = \sum_{c, d \in D_a} r_{ab}^2 \Sigma_p(c, d) + x_{ab}^2 \Sigma_q(c, d) + 2r_{ab}x_{ab}\Sigma_{pq}(c, d)$  **then**
- 7:       Draw edge between nodes  $a$  and  $b^*$
- 8:        $D_{b^*} \leftarrow D_{b^*} \cup D_a$
- 9:        $L \leftarrow L - \{a\}$
- 10:   **end if**
- 11:   **end for**
- 12:    $L \leftarrow L \cup \{b^*\}$
- 13: **end while**

Step 6 adds edges between node  $b^*$  and nodes in set  $L$  of the growing tree using Lemma 3. In the ideal case when infinitely many voltage magnitude samples are collected, second-order moments of the power injections satisfy the relation in Lemma 3. However, we have a finite number of samples. Thus the presence of an edge is determined in Algorithm 1 by checking if the relative difference between the reals on the left and right sides of the condition in Step 6 is less than a predefined tolerance,  $\tau$ :

$$1 - \left| \frac{\sum_{j=1}^m (\epsilon_a^j - \epsilon_{b^*}^j)^2 / m}{\sum_{c, d \in D_a} r_{ab}^2 \Sigma_p(c, d) + x_{ab}^2 \Sigma_q(c, d) + 2r_{ab}x_{ab}\Sigma_{pq}(c, d)} \right| < \tau \quad (23)$$

Steps 9 and 12 update the set of current leaves  $L$  before repeating the reconstruction steps with a new undiscovered node.

**Algorithm Complexity:** Ignoring complexity of computing the second moments in Steps 4 and 6 (part of the data pre-processing), there are  $N + K$  steps ( $N$  load nodes and  $K$  substation nodes) in the ‘while’ loop, and at most  $N$  comparisons in the ‘for’ loop for each node. Therefore, the worst-case complexity of this algorithm is  $O(N^2 + NK)$ .

Note that we use LC-PF model in Algorithm 1. To design the DC-resistive version of Algorithm 1, the condition in Step 6 should be replaced with the result in Lemma 2. (Required modifications are straightforward and thus their description is omitted.)

## VI. EXPERIMENTS

We perform a set of numerical experiments to test and demonstrate the performance of Algorithm 1 in extracting the operational radial forest  $\mathcal{F}$  from meshed “as-designed” distribution networks  $\mathcal{G}$ . We remind the reader that the observer in Algorithm 1 has information of the full graph  $\mathcal{G}$ , the



impedance (resistance and reactance) of all lines (operational or open) as well as the number of connected substation buses. Further, true second moments of active and reactive power injections at each non-substation node are assumed to be known. The set of measurements available as input with the observer comprises of deviations in voltage magnitudes ( $\epsilon$ ) at the grid nodes.

Table II summarizes the distribution grid test systems by the number of load busses, number of substation busses, and the number of tie switches. Additional information on these test systems [23]–[25] can be found online at [26]. In normal operation, each test grid consists of nodes in a forest- or tree-like configuration  $\mathcal{F}$  with open tie-switches. We construct the complete meshed network  $\mathcal{G}$  by closing all the tie-switches. To test the scalability of our algorithms, we increase the number of possible forest configuration by introducing several additional non-operational lines into each system as noted in Table II. Although these contribute to  $\mathcal{G}$ , they are kept open and do not contribute to power flows in the operation forest  $\mathcal{F}$ . The impedances (reactances and resistances) of the additional lines are generated by assigning random values uniformly between the minimum and maximum impedances of the operational lines. Fig. 4 displays the test networks  $\mathcal{G}$  (solid and dashed lines) and the respective operational forests  $\mathcal{F}$  (solid lines only).

TABLE II  
SUMMARY OF THE TESTED DISTRIBUTION GRIDS

Test Case	Number of buses / substations / tie-switches	Non-operational lines added	Source
<i>bus_13_3</i>	13/3/3	10	[23]
<i>bus_29_1</i>	29/1/1	20	[24]
<i>bus_83_11</i>	83/11/13	30	[25]

For each numerical experiment on a grid from Table II, we pick an operational spanning forest  $\mathcal{F}$  by opening tie switches. We also choose the statistics of the injections at each node using Gaussian distributions, unless otherwise specified. These distributions are used to generate multiple power injection samples, and from each vector-valued sample, we solve a PF to compute voltages and phases at every node in the network with the voltages at the substations fixed (i.e. they are slack busses). Averaging over all PF solutions, we compute empirical correlations of voltage magnitudes and phases. Using only these correlations, we run our algorithms and compare the resulting reconstruction with the actual operational configuration. In the reconstruction, we assume the observer has access to the resistance and reactance of all the lines in  $\mathcal{G}$ . All powers and voltages are presented in per unit (p.u.) values.

Figs. 5(a)-5(c) display the accuracy of Algorithm 1 for the different test grids from Table II. The relative error is defined to be the number of mislabeled lines (connected when actually open and vice versa) divided by the size of the operational edge set. The relative error is averaged by computing many reconstructions using the same nodal power injection distributions. Different curves (colors) in Fig. 5(a) show the effect of changing the tolerance  $\tau$  in Eq. 23.

The average fractional error in Figs. 5(a)-5(c) decays ex-

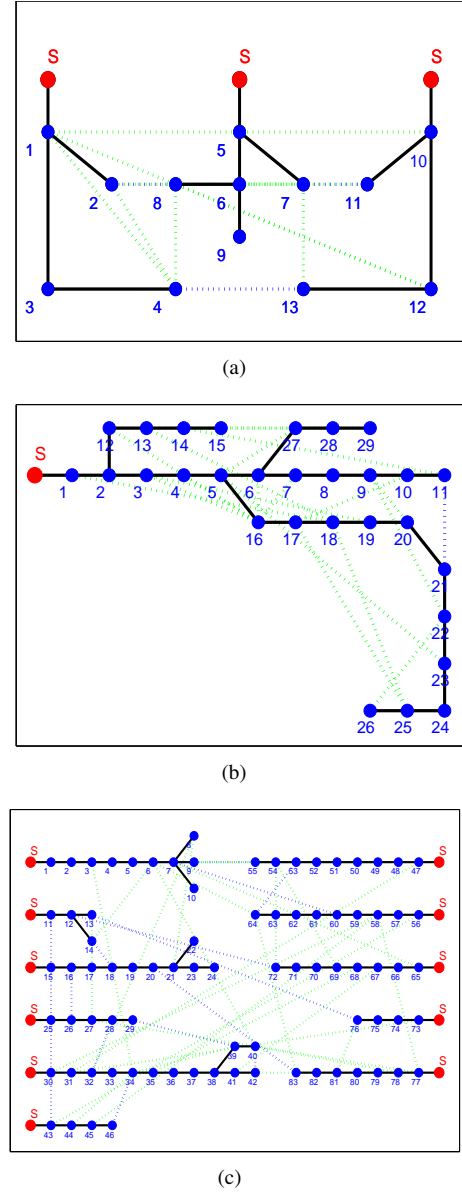


Fig. 4. Layouts of distribution grids, also showing operational cases tested, in accordance with description in the text and summary in Table II. The red circles represent substations (marked as  $S$ ). The blue circles represent load nodes that are numbered. Black lines represent operational edges, while dotted blue lines represent open tie-switches. The additional lines are represented by dotted green lines. (a) *bus\_13\_3* test case (b) *bus\_29\_1* test case (c) *bus\_83\_11* test case

ponentially with the number of samples used by the observer in the learning algorithm. For the tolerance values considered here, the majority of structural errors arise due to connected nodes not satisfying condition (23) and hence being labelled as open. The decay is then intuitive as an increase in the sample size makes empirical moments in voltage magnitudes approximate their true values better which in turn leads to an increase in the number of operational lines satisfying (23) and being correctly identified. On the other hand, if a sufficiently large value of  $\tau$  is used, condition (23) will be relaxed and possibly be satisfied even by unconnected nodes. In such a case, a majority of errors will be recorded due to open lines being incorrectly labelled as operational. As errors of this type (open edges classified as operational) does not improve with

the number of measurement samples, the average fractional errors will not decay with the sample size. This is elucidated in Fig. 6, where the *bus\_13\_3* structure is learnt, in the presence of 50 non-operational lines. Note that for larger values of  $\tau$  in Fig. 6, the errors do not decay with the sample size whereas for smaller values, they do as justified in the preceding discussion. For larger sample sizes, a smaller value of  $\tau$  is indeed preferable as observed in Fig. 6.

## VII. CONCLUSIONS & PATH FORWARD

Accurate structural estimation of distribution grids is important to many applications including failure identification, power flow optimization and estimation of state variables. In this manuscript, we have developed algorithms to learn the structure of a radial distribution grid using observed nodal voltage magnitude measurements. We have used a Linear Coupled (LC) approximation that relates complex nodal voltages to the complex power consumptions to prove that second moments of nodal voltage magnitudes in radial distribution grids follow certain statistical structure/ordering. Our algorithm relies on these results to reconstruct the operational tree in a bottom up fashion – starting from the leaves and progressing to the root of the grid.

The primary benefits of our approach are two-fold. First, our model is practical as voltage measurements are easily available at distribution grid nodes and individual devices. Second, the only assumption used regarding the statistics of the loads is the positivity of non-central correlation of nodal load profiles, which is natural for most distribution grids and more general than other assumptions discussed in the literature. We tested our algorithms on sample distribution grids and observed an exponential decay of average errors with increasing number of measurement samples. In the next part, we extend the work in this paper along the following directions: coupling structure learning with estimation of load statistics in the grid or estimation of line parameters in the grid, and learning grid structure even when measurement data is missing. Developing general algorithms for learning distribution grid operational forests that are not restricted to linearized power flow models remains a potential future direction of research.

## APPENDIX

### A. Basic Power Flows

Note that the operational ‘base-constrained spanning forest’  $\mathcal{F}$  can be thought of as a spanning tree over an extended graph, where an (artificial) super node is introduced and connected with (artificial) lines to all the sub-station nodes. This trick allows, without a loss of generality, to limit our discussion in the following to spanning trees, thus replacing  $\mathcal{F}$  by  $\mathcal{T}$ .

Let  $z = (z_{ab} = r_{ab} + ix_{ab} | (ab) \in \mathcal{E})$  as the vector of complex line impedances in the grid. ( $i^2 = -1$ .) Expressed in terms of the complex powers and potentials (voltages and phases) the Kirchoff laws over the operational (spanning tree) configuration  $\mathcal{T}$  become

$$\forall a \in \mathcal{V} : P_a = p_a + iq_a = \sum_{b:(ab) \in \mathcal{E}^{\mathcal{T}}} \frac{v_a^2 - v_a v_b \exp(i\theta_a - i\theta_b)}{z_{ab}^*} \quad (24)$$

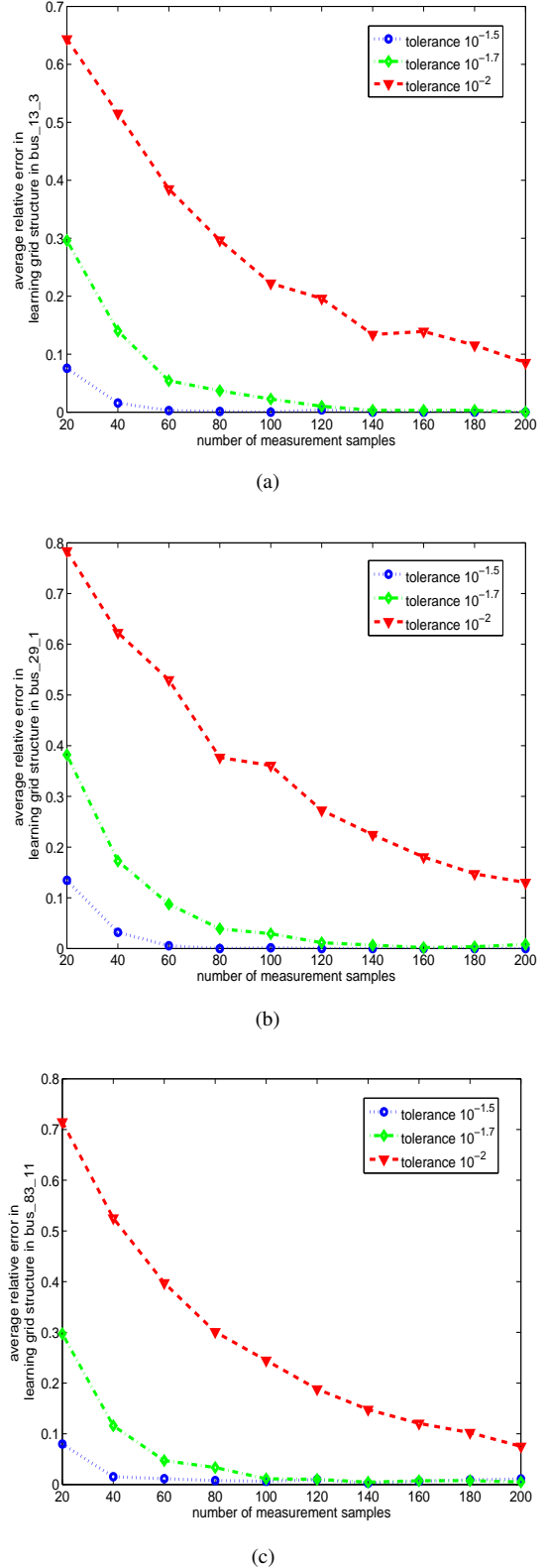


Fig. 5. Average fractional errors plotted against the number of measurement samples for the LC-PF model with different tolerance thresholds - correspondent to the Algorithm 2.(a), (b) and (c) show results for the *bus\_13\_3*, *bus\_29\_1* and *bus\_83\_11* models respectively.

where the real valued scalars,  $v_a$  and  $\theta_a$ , characterize voltage magnitude and phase respectively at node  $a$ . We assume that the power within the system considered is balanced through



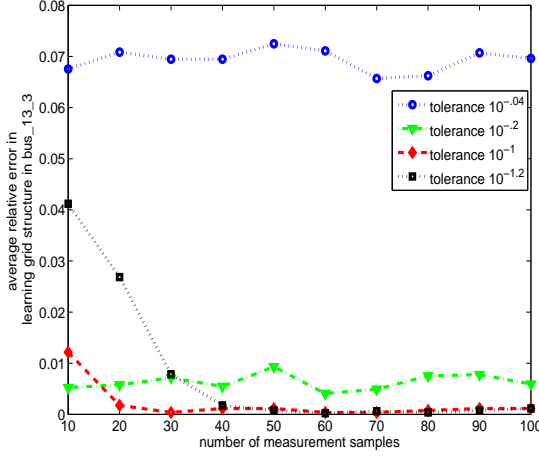


Fig. 6. Effect of threshold on the average fractional errors in Algorithm 1 using the LC-PF model for the *bus\_13\_3* (with 50 additional non-operational edges).

a slack bus,  $a = 0$ . The set of Eqs. (24), expressing potentials via complex powers injected at the nodes of the power graph, are called *Power Flow* (PF) equations.

### B. Linear Coupled (LC) Approximation of Power Flows

We linearize the PF Eqs. (24) in the first order jointly over phase difference and voltage deviations ( $v_a - 1 = \epsilon_a$ ) from nominal, i.e. the two corrections are considered on equal footing. We arrive at the following set of equations:

$$p_a = \sum_{b:(ab) \in \mathcal{E}^T} (\beta_{ab}(\theta_a - \theta_b) + g_{ab}(\epsilon_a - \epsilon_b)), \quad (25)$$

$$q_a = \sum_{b:(ab) \in \mathcal{E}^T} (-g_{ab}(\theta_a - \theta_b) + \beta_{ab}(\epsilon_a - \epsilon_b)) \quad (26)$$

$$\text{where } \forall a \in \mathcal{V}, \forall (ab) \in \mathcal{E}^T: |\epsilon_a| \ll 1, |\theta_a - \theta_b| \ll 1, \quad (27)$$

$$g_{ab} \doteq \frac{r_{ab}}{x_{ab}^2 + r_{ab}^2}, \beta_{ab} \doteq \frac{x_{ab}}{x_{ab}^2 + r_{ab}^2} \quad (28)$$

Eqs. (25,26) show coupling between phases and voltages, thus calling it the Linear-Coupled (LC) approximation, is proper. The linearity of LC-PF is used in the paper in deriving results to learn the grid structure. Two comments are in order. First of all, notice that the LC-PF approximation does not make any assumption about the relative strength of inductance and resistance, thus making it applicable to power distribution systems where the two line characteristics are typically of the same order. Second, expressions under the sum on the rhs of Eqs. (25,26) represent active and reactive power flows which are antisymmetric ( $p_{a \rightarrow b} = -p_{b \rightarrow a}$ ,  $q_{a \rightarrow b} = -q_{b \rightarrow a}$ ). This emphasizes an important consequence of linear approximation in the LC-PF model – both active and reactive losses in lines are ignored as such losses occur at second order.

### C. DC-resistive approximation

In low-voltage distribution grids line inductances may be much smaller in magnitude than line resistances. Then the LC-PF model can be simplified even further. Indeed, taking this case to the extreme where inductance can be ignored

in comparison with the resistance,  $\forall \{a, b\}: \beta_{ab} \ll g_{ab}$ , we arrive at the following resistance-dominating version of Eqs. (25,26)

$$\forall a \in \mathcal{V}: p_a \approx \sum_{b:(ab) \in \mathcal{E}^T} g_{ab}(\epsilon_a - \epsilon_b), q_a \approx \sum_{b:(ab) \in \mathcal{E}^T} g_{ab}(\theta_b - \theta_a)$$

We will coin this approximation DC-resistive PF. Its formulation is similar to the traditional DC flow model [27] where active power flows are related to phase angles. The traditional DC model, used primarily for transmission networks, requires line inductances to dominate resistances, which is seldom observed in distribution grids.

### D. From Power Flows to DistFlow and LinDistFlow

The DistFlow Eqs. introduced by Baran and Wu in [17], [18] are derived from the PF Eq. (24)

$$p_{a \rightarrow b} - r_{ab} \frac{p_{a \rightarrow b}^2 + q_{a \rightarrow b}^2}{v_a^2} = p_b + \sum_{(bc) \in \mathcal{E}^T; c \neq a} p_{b \rightarrow c}, \quad (29)$$

$$q_{a \rightarrow b} - x_{ab} \frac{p_{a \rightarrow b}^2 + q_{a \rightarrow b}^2}{v_a^2} = q_b + \sum_{(bc) \in \mathcal{E}^T; c \neq a} q_{b \rightarrow c}, \quad (30)$$

$$v_b^2 = v_a^2 - 2(r_{ab}p_{a \rightarrow b} + x_{ab}q_{a \rightarrow b}) + (r_{ab}^2 + x_{ab}^2) \frac{p_{a \rightarrow b}^2 + q_{a \rightarrow b}^2}{v_a^2} \quad (31)$$

where each of the three equations above are stated in terms of active,  $p_{a \rightarrow b}$ , and reactive,  $q_{a \rightarrow b}$ , powers over directed lines,  $a \rightarrow b$ , and voltages over nodes,  $a \in \mathcal{V}$ . If power losses at any line segment is negligible, Eqs. (29,30,31) reduce to [4].

$$p_{a \rightarrow b} \approx p_b + \sum_{(bc) \in \mathcal{E}^T; c \neq a} p_{b \rightarrow c}, q_{a \rightarrow b} \approx q_b + \sum_{(bc) \in \mathcal{E}^T; c \neq a} q_{b \rightarrow c}, \quad (32)$$

$$\phi_b \approx \phi_a - 2(r_{ab}p_{a \rightarrow b} + x_{ab}q_{a \rightarrow b}), \quad \phi_a \equiv v_a^2 \quad (33)$$

On a general graph with loops, the number of these (directed) edge-related variables in Eq. (24) is larger than the number of phases and voltages. However, the former becomes equal to the later if the grid is a tree, where the number of edges is equal to the number of nodes minus one. Therefore when the phase and voltage at one special node (each substation node per tree in our case) is fixed at  $v_0$  (say), phases and voltages at all other nodes can be reconstructed from the directional line flows, and vice versa.

Note that additional boundary conditions arise in a tree due to the requirement of active and reactive power flowing from/into any leaf being equal to its nodal injection. Thus,

$$\forall a \in \mathcal{V}_0: v_a = v_0, \quad (34)$$

$$\forall b \in \mathcal{V}_l^T, \& (a \rightarrow b) \in \mathcal{E}^T: p_{a \rightarrow b} = p_b, \quad q_{a \rightarrow b} = q_b, \quad (35)$$

where  $\mathcal{V}_0 \subset \mathcal{V}$  is the set of  $K$  substations nodes (colored in red in Fig. 1), and  $\mathcal{V}_l^T$  is the set of leaf-nodes in tree  $\mathcal{T}$ . Based on the assumption that voltage drop across any line segment is sufficiently small, we linearize Eq. (33) by substituting  $v_a = 1 + \epsilon_a$ , with  $|\epsilon_a| \ll 1$  and have

$$\forall (ab) \in \mathcal{E}^T: \epsilon_a - \epsilon_b \approx (r_{ab}p_{a \rightarrow b} + x_{ab}q_{a \rightarrow b}). \quad (36)$$

Notice that the LinDistFlow Eqs. (32,36), derived assuming that the grid graph is a tree, are exactly equivalent to Eqs. (25,26) of the Linear Coupled (LC) Approximation.

#### ACKNOWLEDGMENT

The work at LANL was funded by the Advanced Grid Modeling Program in the Office of Electricity in the US Department of Energy and was carried out under the auspices of the National Nuclear Security Administration of the U.S. Department of Energy at Los Alamos National Laboratory under Contract No. DE-AC52-06NA25396.

#### REFERENCES

- [1] R. Hoffman, "Practical state estimation for electric distribution networks," in *Power Systems Conference and Exposition, 2006. PSCE'06. 2006 IEEE PES*. IEEE, 2006, pp. 510–517.
- [2] J. Liu, Y. Xiao, S. Li, W. Liang, and C. Chen, "Cyber security and privacy issues in smart grids," *Communications Surveys & Tutorials, IEEE*, vol. 14, no. 4, pp. 981–997, 2012.
- [3] Y. Sharon, A. M. Annaswamy, A. L. Motto, and A. Chakraborty, "Topology identification in distribution network with limited measurements," in *Innovative Smart Grid Technologies (ISGT), 2012 IEEE PES*. IEEE, 2012, pp. 1–6.
- [4] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *Power Delivery, IEEE Transactions on*, vol. 4, no. 2, pp. 1401–1407, 1989.
- [5] J. A. P. Lopes, F. J. Soares, and P. M. R. Almeida, "Integration of electric vehicles in the electric power system," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 168–183, 2011.
- [6] K. Turitsyn, P. Sulc, S. Backhaus, and M. Chertkov, "Options for control of reactive power by distributed photovoltaic generators," *Proceedings of the IEEE*, vol. 99, no. 6, pp. 1063–1073, 2011.
- [7] J. Kim and L. Tong, "On topology attack of a smart grid," in *Innovative Smart Grid Technologies (ISGT), 2013 IEEE PES*. IEEE, 2013, pp. 1–6.
- [8] D. Deka, R. Baldick, and S. Vishwanath, "Attacking power grids with secure meters: The case for breakers and jammers," in *Computer Communications Workshops (INFOCOM WKSHPS), 2014 IEEE Conference on*. IEEE, 2014, pp. 646–651.
- [9] M. J. Wainwright and M. I. Jordan, "Graphical models, exponential families, and variational inference," *Foundations and Trends® in Machine Learning*, vol. 1, no. 1-2, pp. 1–305, 2008.
- [10] P. Ravikumar, M. J. Wainwright, J. D. Lafferty *et al.*, "High-dimensional ising model selection using 1-regularized logistic regression," *The Annals of Statistics*, vol. 38, no. 3, pp. 1287–1319, 2010.
- [11] A. Anandkumar, V. Tan, and A. S. Willsky, "High-dimensional graphical model selection: tractable graph families and necessary conditions," in *Advances in Neural Information Processing Systems*, 2011, pp. 1863–1871.
- [12] P. Netrapalli, S. Banerjee, S. Sanghavi, and S. Shakkottai, "Greedy learning of markov network structure," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*. IEEE, 2010, pp. 1295–1302.
- [13] P. Ravikumar, M. J. Wainwright, G. Raskutti, B. Yu *et al.*, "High-dimensional covariance estimation by minimizing 1-penalized log-determinant divergence," *Electronic Journal of Statistics*, vol. 5, pp. 935–980, 2011.
- [14] V. Kekatos, G. B. Giannakis, and R. Baldick, "Grid topology identification using electricity prices," *arXiv preprint arXiv:1312.0516*, 2013.
- [15] M. He and J. Zhang, "A dependency graph approach for fault detection and localization towards secure smart grid," *Smart Grid, IEEE Transactions on*, vol. 2, no. 2, pp. 342–351, 2011.
- [16] S. Bolognani, N. Bof, D. Michelotti, R. Muraro, and L. Schenato, "Identification of power distribution network topology via voltage correlation analysis," in *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*. IEEE, 2013, pp. 1659–1664.
- [17] M. Baran and F. Wu, "Optimal sizing of capacitors placed on a radial distribution system," *Power Delivery, IEEE Transactions on*, vol. 4, no. 1, pp. 735–743, Jan 1989.
- [18] —, "Optimal capacitor placement on radial distribution systems," *Power Delivery, IEEE Transactions on*, vol. 4, no. 1, pp. 725–734, Jan 1989.
- [19] <http://www.powerstandards.com/PQube.php>.
- [20] A. Phadke, "Synchronized phasor measurements in power systems," *Computer*, 1993.
- [21] J. Resh, "The inverse of a nonsingular submatrix of an incident matrix," *IEEE Transactions on Circuit Theory*, vol. 10, p. 131.
- [22] L. Pedersen, J. Stang, and R. Ulseth, "Load prediction method for heat and electricity demand in buildings for the purpose of planning for mixed energy distribution systems," *Energy and Buildings*, vol. 40, no. 7, pp. 1124–1134, 2008.
- [23] S. Civanlar, J. Grainger, H. Yin, and S. Lee, "Distribution feeder reconfiguration for loss reduction," *Power Delivery, IEEE Transactions on*, vol. 3, no. 3, pp. 1217–1223, Jul 1988.
- [24] U. Eminoglu and M. H. Hocaoglu, "A new power flow method for radial distribution systems including voltage dependent load models," *Electric Power Systems Research*, vol. 76, no. 13, pp. 106 – 114, 2005.
- [25] C.-T. Su, C.-F. Chang, and J.-P. Chiou, "Distribution network reconfiguration for loss reduction by ant colony search algorithm," *Electric Power Systems Research*, vol. 75, no. 23, pp. 190 – 199, 2005.
- [26] [Online]. Available: <http://www.dejazzzer.com/reds.html>
- [27] A. Abur and A. G. Exposito, *Power system state estimation: theory and implementation*. CRC Press, 2004.



**Deepjyoti Deka** Deepjyoti Deka received his M.S. in Electrical Engineering from University of Texas, Austin in 2011, and his B.Tech in Electronics and Communication Engineering from IIT Guwahati, India, in 2009 for which he was awarded the Institute Silver Medal. He is currently a PhD candidate in Electrical Engineering at UT Austin. His research focusses on the design and analysis of power grid structure, operations and data security. He is also interested in modeling and optimization in social and physical networks. He has held internship positions at Los Alamos National Lab, Los Alamos NM, Electric Reliability Council of Texas, Taylor TX, and Qualcomm Inc, San Diego CA.



**Scott Backhaus** Scott Backhaus received the Ph.D. degree in physics from the University of California at Berkeley in 1997 in the area of experimental macroscopic quantum behavior of superfluid He-3 and He-4. In 1998, he came to Los Alamos, NM, was Director's Funded Postdoctoral Researcher from 1998 to 2000, a Reines Postdoctoral Fellow from 2001 to 2003, and a Technical Staff Member from 2003 to the present. While at Los Alamos, he has performed both experimental and theoretical research in the area of thermoacoustic energy conversion for which he received an R&D 100 award in 1999 and Technology Review's Top 100 Innovators Under 35 [award in 2003]. Recently, his attention has shifted to other energy-related topics including the fundamental science of geologic carbon sequestration and grid-integration of renewable generation.



**Michael Chertkov** Dr. Chertkov's areas of interest include statistical and mathematical physics applied to energy and communication networks, machine learning, control theory, information theory, computer science, fluid mechanics and optics. Dr. Chertkov received his Ph.D. in physics from the Weizmann Institute of Science in 1996, and his M.Sc. in physics from Novosibirsk State University in 1990. After his Ph.D., Dr. Chertkov spent three years at Princeton University as a R.H. Dicke Fellow in the Department of Physics. He joined Los Alamos National Lab in 1999, initially as a J.R. Oppenheimer Fellow in the Theoretical Division. He is now a technical staff member in the same division. Dr. Chertkov has published more than 130 papers in these research areas. He is an editor of the Journal of Statistical Mechanics (JSTAT), associate editor of IEEE Transactions on Control of Network Systems, a fellow of the American Physical Society (APS), and a Founding Faculty Fellow of Skoltech (Moscow, Russia).